IMPROVED STRATEGIES FOR MTDNA SEQUENCE ANALYSIS OF HIGHLY DEGRADED FORENSIC REMAINS

Matthew N. Gabriel, M.F.S., Edwin F. Huffine, M.S., John H. Ryan, Ph.D., Christopher W. Los, M.S.F.S., Nicholas C.S. Yang, M.F.S., Mitchell M. Holland, Ph.D., and Thomas J. Parsons, Ph.D.

The Armed Forces DNA Identification Laboratory, United States Department of Defense, United States Department of the Army

INTRODUCTION

Mitochondrial DNA (mtDNA) analysis of highly degraded skeletal remains is often used for forensic identification due largely to the high genome copy number per cell. Literature from the 'ancient DNA' field has shown that highly degraded samples contain populations of intact DNA molecules that are severely restricted in size (1-4). Handt *et al.* have demonstrated the targeting and preferential amplification of authentic human DNA sequences with small amplicon products of 150bp or less (1,2). Given this understanding of ancient DNA preservation and amplification, we have completed a validation study demonstrating an improved approach to mtDNA analysis of hypervariable regions 1 and 2 (HV1/HV2) in the most highly degraded forensic specimens (manuscript in preparation). This 'mini-primer set' (MPS) amplification strategy consists of four overlapping products that span each of the HV regions and range from 126-170bp, with an average size of 141bp (Figure 1). We have demonstrated a dramatic increase in amplification success with these primer pairs when compared to data obtained using larger ~250bp amplicons. Further, by targeting the abundant population of molecules in this size range, samples previously 'unreported' due to mixed sequences provided a single, authentic sequence using the MPS method.

This strategy is suitable for AFDIL purposes as difficult cases often entail analysis of extremely degraded samples that have been exposed to soil and elevated temperatures for decades. In some instances, samples are in such poor condition that multiple amplification attempts with ~250bp primer sets provide no product. In addition to the limited availability of intact molecules present in these samples, large amounts of bacterial and fungal DNA relative to authentic template DNA complicate ancient remains analysis. The high sequence complexity caused by multiple distinct genomes challenges PCR specificity greatly and often leads to the production of nonspecific minor or predominant artifacts and amplicon smearing. Three samples in the initial validation study produced nonspecific artifacts following amplification with MPS2B (F16268/R16410). For these samples, standard purification methods based on molecular weight cutoffs were not effective in isolating the desired product from other nonspecific products. We report here a rapid purification method for the isolation of desired PCR products from nonspecific artifacts that uses ion-pairing reverse-phase high performance liquid chromatography (IPRP-HPLC) coupled with the DNASep technology offered by the Transgenomic Wave DNA Fragment Analysis System™. This purification method also offers increased sequence data quality in routine cases and provides accurate DNA quantitation as well as integration with high throughput platforms. The MPS amplification strategy then, used in combination with alternative purification methods such as IPRP-HPLC, greatly expands the range of degraded forensic specimens available for mtDNA testing and moves toward greater levels of sensitivity.

MATERIALS AND METHODS

DNA extraction

DNA was extracted from dried skeletal remains using standard AFDIL protocols. The initial weights of the bone specimens submitted ranged from 1.2-4.3g. The outer and inner surfaces were sanded thoroughly to remove spongy bone material and the resulting fragments were measured at 0.4-0.9g. The samples were washed thoroughly with distilled water and absolute ethanol and each specimen was then pulverized into a fine powder. The powders were then incubated overnight at 56°C in the presence of

extraction buffer (10mM Tris, pH 8.0, 100mM NaCl, 50mM EDTA (pH 8.0), 0.5% SDS + 0.1mL of 20mg/mL proteinase K). Phenol:chloroform:isoamoyl alcohol extraction was performed with an equal volume, and separation of aqueous and organic layers was achieved by centrifugation. The upper layer was recovered and the extraction repeated several times. An n-butanol extraction was performed, and the lower aqueous layer was recovered, added to a Centricon-100 (Amicon, Inc., Beverly, MA) concentrator, and centrifuged. The filtrate was discarded, and two Tris-EDTA buffer washes were performed prior to sample recovery.

MPS amplification with non-probative extracts

PCR reactions were carried out in 50µL volumes using 1X AmpliTaq® polymerase buffer with 1.5mM $MgCl_2$ final concentration (*PE* Applied Biosystems, Foster City, CA), 200µM each dNTP (Boehringer Manheim, Indianapolis, IN), 0.025mg/mL bovine serum albumen (Life Technologies, Gaithersburg, MD), 0.4µM each forward and reverse amplification primer (synthesized in-house with ABI 392 DNA Synthesizer), and 5 units of AmpliTaqGold® polymerase (*PE* Applied Biosystems). All reactions were performed using an initial 96.0C activation step for 10 minutes with twenty-second strand separation and primer annealing steps (96.0C and 48.0C, respectively), followed by a thirty-second extension step (72.0C) for 42 cycles. Non-probative extracts were amplified with similar amounts of extract (2.0µL) and AmpliTaqGold® polymerase (12.5 units) as used for previous primer set amplifications of ~250bp. For each group of samples, a positive control amplification with 100pg of high quality genomic DNA of known mtDNA sequence was included along with two negative control amplifications and the corresponding reagent blanks to monitor contamination levels. In each case, 10% of the amplification reaction was analyzed by agarose gel electrophoresis stained with ethidium bromide to note DNA band intensity and resolution.

PCR product purification

Products were purified using either the standard molecular weight cutoff offered by Centricon-30 (Amicon, Inc., Beverly, MA) concentrators or the IPRP-HPLC purification with the DNASep technology of the Transgenomic Wave System™ (Transgenomic, Inc., Omaha, NE). For the Centricon-30 purification, the amplification product was added to 2mL deionized water and centrifugation was performed for 30 minutes at 4000rpm. A second 2mL water wash was performed as before and the retentate was recovered through centrifugation. For IPRP-HPLC purification, 5µL of the amplification product was initially injected onto the DNASep column to observe the approximate retention time and chromatographic output for each sample. Table 1 describes the method used for sample injection and fragment elution. A determination was then made for peak collection parameters using the Transgenomic Fragment Collector™ (Transgenomic, Inc., Omaha, NE) and a second 5.0µL injection was performed. The collection parameters were as follows: Absolute threshold was set at 0.8mV using the 'peak' collection option. Beginning and ending collection times were set at 6.0 and 6.5min, respectively, and the total collection time of 0.5min was entered. Following collection, the ~150µL volumes were dried down in a speed vacuum and resuspended in 10µL deionized water. 1.0µL was then used for each cycle sequencing reaction.

Cycle sequencing

Approximately 5-20ng of Centricon-purified product was used for cycle sequencing with BigDye Terminator chemistry (*PE* Applied Biosystems). Cycle sequencing was done in 20µL volumes with 8.0µL of BigDye Terminator mix and 0.5µM of either the forward or reverse primer as used for amplification. Cycle sequencing was performed as follows: 96.0C for 15sec, 50.0C for 5sec, 60.0C for 4.0min X 25 cycles. One positive sequence control was performed for each group of samples with pGEM plasmid and M13 primer. Sequencing reactions were purified with Centriflex AGTC Gel Filtration Cartridges (Edge BioSystems, Gaithersburg, MD) and dried in a speed vacuum. Samples were resuspended in 4.0µL of loading buffer (deionized formamide + 50mg/mL blue dextran in 25mM EDTA in a 5:1 ratio), and a standard 4.0% polyacrylamide gel (19:1) was run on the ABI 377 DNA Sequencer. Data collection and analysis was performed with Sequence ™ software (Applied Biosystems).

RESULTS AND DISCUSSION

Validation of the MPS amplification strategy clearly demonstrates the usefulness of shorter primer pairs to effectively target the high abundance population of intact molecules present in highly degraded remains. On a subset of particularly difficult samples chosen for this study, the MPS strategy exhibits an increase in amplification success when compared to larger ~250bp amplicons. This same subset of difficult samples in some instances also produced an apparent mixture of two sequences under standard primer set testing, rendering them unreportable. These mixtures result from co-amplification of low level contaminating sequences. The use of MPS primers on these samples in most cases resulted in recovery of only one of the two mixed sequences, presumably by targeting a proportionally more abundant population of short authentic template molecules (manuscript in preparation). While results of testing with standard primer sets remain robust, the MPS strategy provides an alternative that can increase the probability of determining reportable sequences in cases of highly degraded samples.

The highly degraded samples routinely tested in our laboratory often contain large amounts of bacterial and fungal DNA relative to authentic human DNA sequences. This sequence complexity challenges PCR specificity greatly, often resulting in the production of either minor or major nonspecific artifacts. For the initial amplification of sample extracts 04A-06A with larger ~250bp primer sets, no sequence data was reported (data not shown). The corresponding MPS amplifications demonstrate the difficulty faced with PCR specificity in aged remains by the production of nonspecific artifact bands. Figure 2 shows the presence of multiple nonspecific artifacts slightly larger than the desired 143bp MPS2B amplicon (F16268/R16410) by agarose gel electrophoresis.

IPRP-HPLC offers an alternative purification method that allows rapid isolation of desired products from nonspecific artifacts. The DNASep technology offered by the Wave System™ enables accurate binding of DNA molecules to the column matrix and elution of DNA fragments based on size (6,7) (Figure 3). In our experience, this technology has the ability to resolve fragments that differ by as little as 7-10bp (data not shown). Figure 4 shows the chromatograms for the corresponding 5.0μL injections of samples 04A, 05A and 06A. In each case, the 143bp fragment exhibits similar retention times (R_T) and provides an accurate representation of the DNA products present in each sample. The data output from the Wave System™ clearly identifies several minor artifacts as definable peaks that are visualized as smears by agarose gel electrophoresis or are so faint that they cannot be seen. This observation demonstrates the greater detection limits offered by this analysis system. After the initial 5.0μL injection to observe the retention times and morphology of the sample chromatographs, peak capture parameters were determined for the isolation of the desired fragment. Absolute threshold values and collection windows were specified to ensure capture of only the 143bp product and not neighboring fragments of similar size. After collection and drying down of the sample recovery, direct sequencing methods provided clean, high quality data. Figure 5 demonstrates the sequence data attainable with this purification method and the corresponding poor quality data provided by the standard Centricon-30 purification method. In addition to fragment isolation, the Wave System™ also has the capability of quantitating PCR product prior to direct sequencing methods to provide accurate and reproducible sequence peak intensity.

A significant drawback of short primer sets is the susceptibility to high sequence background levels. In our experience, sequence background due to poor template quality often can be lessened by using internal sequencing primers. Due to the small amplicon size of the short primer sets, internal sequencing primers decrease the total available sequence information that can be gained with each sequence run. This strategy then lessens the benefit of MPS amplification and therefore is not a feasible option for complete sequence reporting. Further, since neighboring MPS amplicons contain short sequence overlaps (in some cases only several bases), ‘late starts’ in sequence reads decrease the available sequence information and in some instances may not provide confirmation between neighboring products. Preliminary data in our laboratory demonstrate the ability of IPRP-HPLC to significantly decrease background levels and prevent sequence read ‘late starts’, thereby providing maximum benefit of this amplification strategy.

REFERENCES

(1) Handt O., Krings M., Ward R.H., Paabo S. The retrieval of ancient human DNA sequences. Am. J. Hum. Genet. 1996;59:368-76.

(2) Handt O., Richards M., Trommsdorff M., Kilger C., Simanainen J., Georgiev O., et al. Molecular genetic analyses of the tyrolean ice man. Science 1994 June;264:1775-8.

(3) Handt O., Hoss M., Krings M., Paabo S. Ancient DNA: methodological challenges. Experientia 1994;50:524-9.

(4) Paabo S. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc. Natl. Acad. Sci. 1989 March;86:1939-43.

(5) Anderson S., Bankier A.T., Barrell B.G., De Bruijn M.H.L., Coulson, A.R., Drouin J., et al. Sequence and organization of the human mitochondrial genome. Nature 1981 April;290:457-65.

(6) Oefner P.J. and Bonn G.K. High-resolution liquid chromatography of nucleic acids. American Laboratory 1994 June; 28C-28J.

(7) Huber C.G., Oefner P.J., Preuss E., and Bonn G.K. High-resolution liquid chromatography of DNA fragments on non-porous poly(styrene-divinylbenzene) particles. Nucleic Acids Research 1993;21:1061-1066.

Figure 1A - Current AFDIL ‘Primer Set’ Amplification Strategy for Degraded Remains

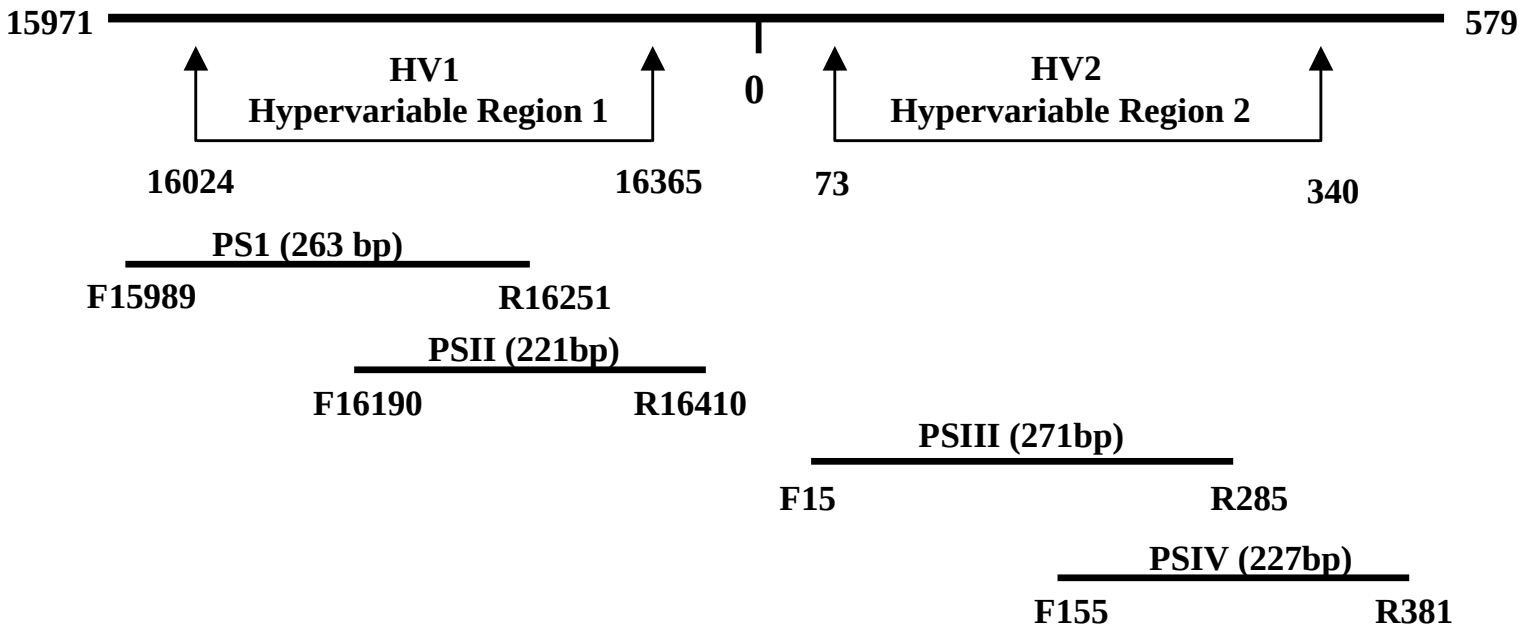

Figure 1A - The current AFDIL ‘primer set’ amplification strategy for highly degraded specimens entails the amplification of two overlapping products for each hypervariable region. The amplicons range in size from 221-271bp and are on average ~250bp. Base calling orientation follows the numbering scheme developed by Anderson *et al.* (5).

Figure 1B - HV1 / HV2 ‘Mini-primer set’ Amplification Strategy

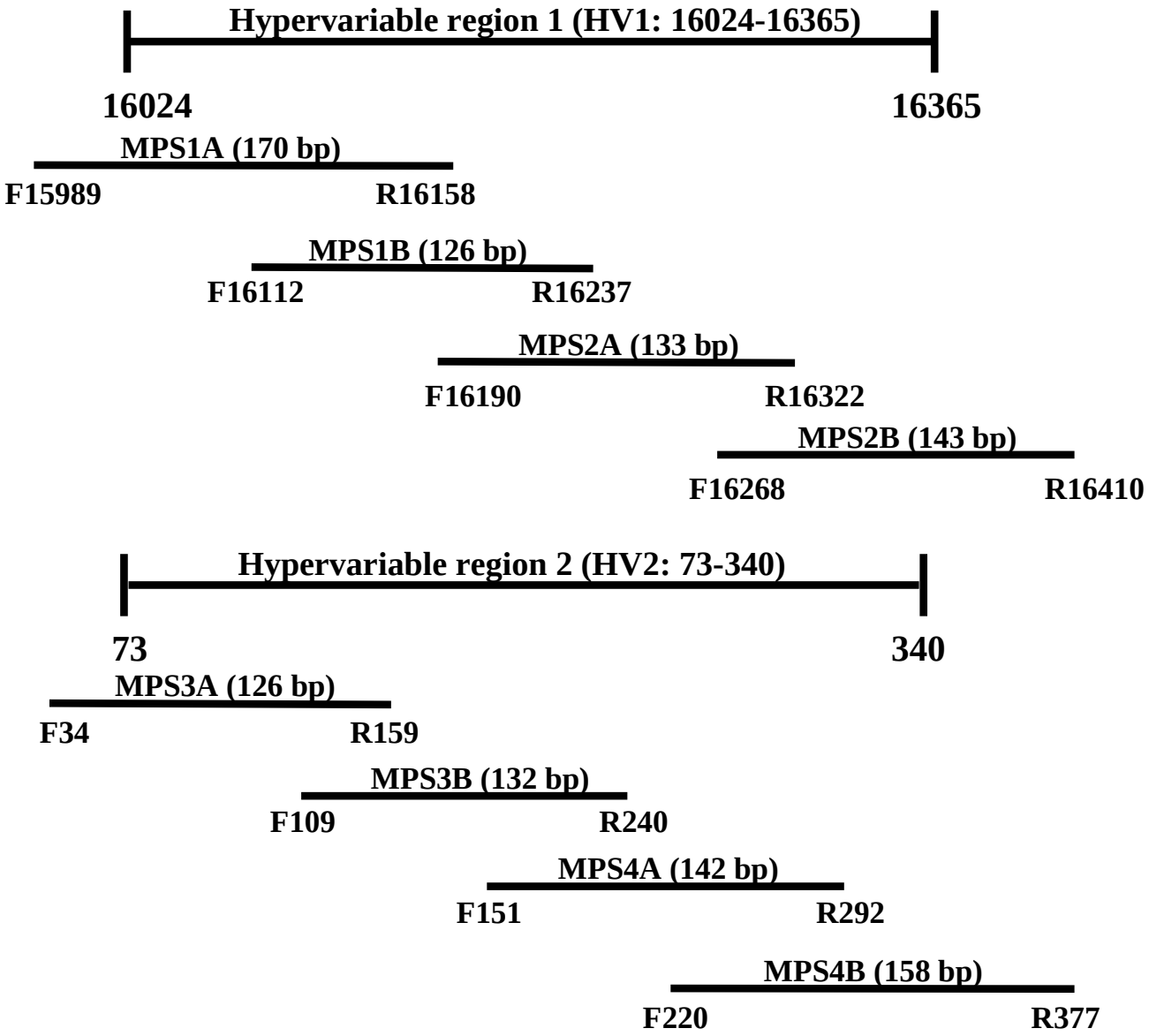

Figure 1B - The ‘Mini-primer set’ amplification strategy for highly degraded specimens involves four overlapping products for each hypervariable region. MPS amplicons range in size from 126-170bp and are on average ~141bp. For HV1 trials, primer pairs were selected from previously published ancient DNA studies (1, Anne Stone dissertation). HV2 primer pairs were designed using either Lasergene PrimerSelect Version 4.0 (DNASTAR, Inc., Madison, WI) or Oligo Version 5.0 (National Biosciences Inc., Plymouth, MN) software programs to determine suitable primer characteristics and “optimal” annealing temperatures. Base calling orientation follows the numbering scheme developed by Anderson *et al.* (5).

Figure 2 - Production of Nonspecific Artifacts with MPS2B Amplification

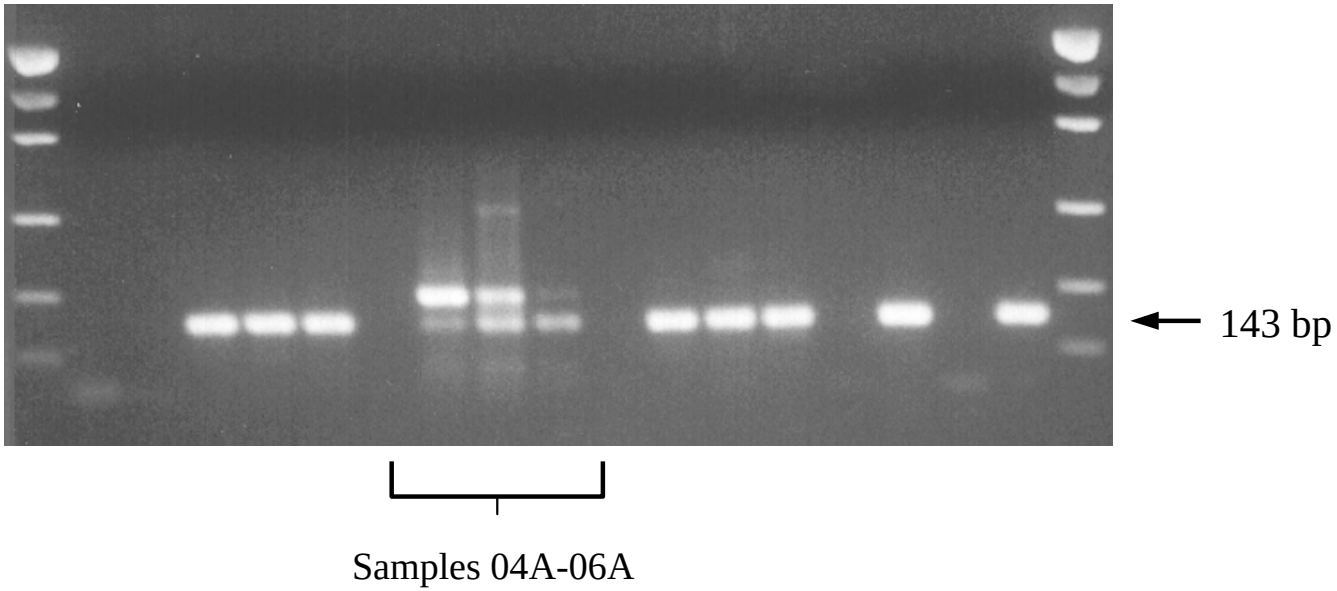

Figure 2 - Agarose gel showing amplification of ten bone extracts with MPS2B (F16268/R16410). Lanes 1 and 19; Low DNA mass ladder. Lanes 8, 9 and 10; Samples 04A1, 05A1 and 06A1 (5.0µL) . In each case (lanes 8-10), nonspecific artifact band production complicates sequence analysis. Alternative purification methods are necessary to provide reportable data quality by direct sequencing methods.

Figure 3 - Ion-pairing Reverse Phase High Performance Liquid Chromatography (IPRP-HPLC)

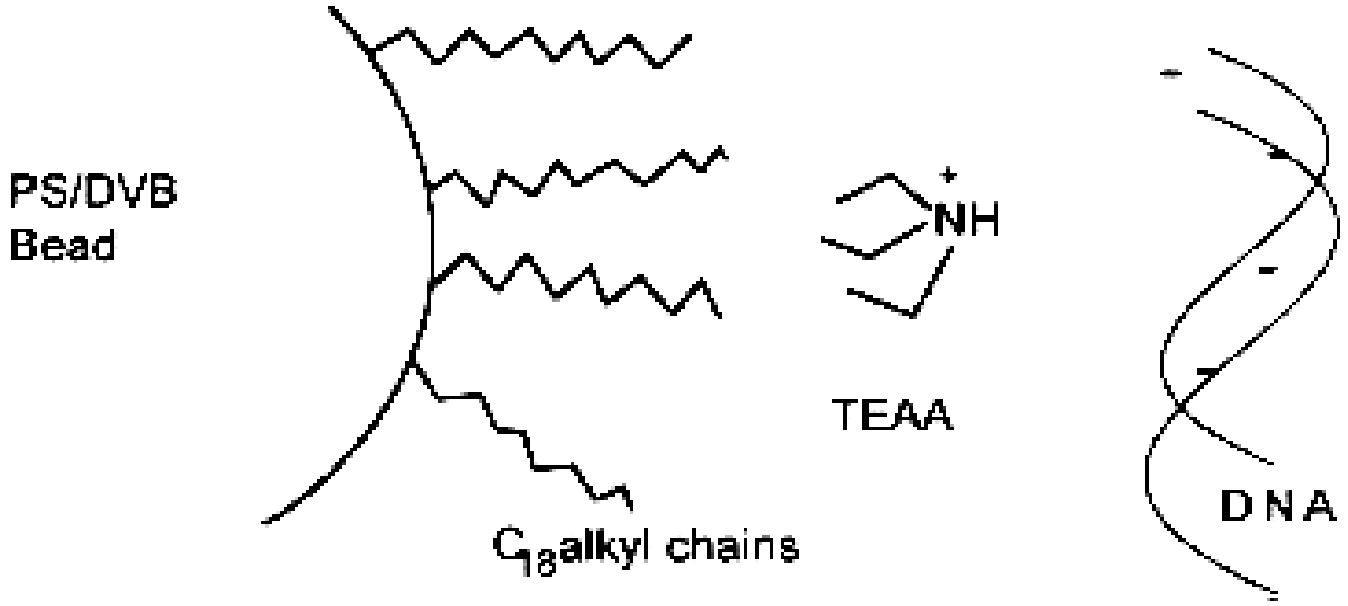

Figure 3 - Ion-pairing reverse-phase high performance liquid chromatography with the Transgenomic Wave System™ uses triethylammonium acetate as a 'bridge' molecule to bind the DNA to the neutral, hydrophobic polystyrene divinyl benzene column matrix. Elution occurs on the basis of fragment size via the introduction of acetonitrile, and the resulting fragment concentration is measured by ultraviolet absorbance at OD_{260}.

Figure 4 - Chromatograms and Data Output for Samples 04A-06A

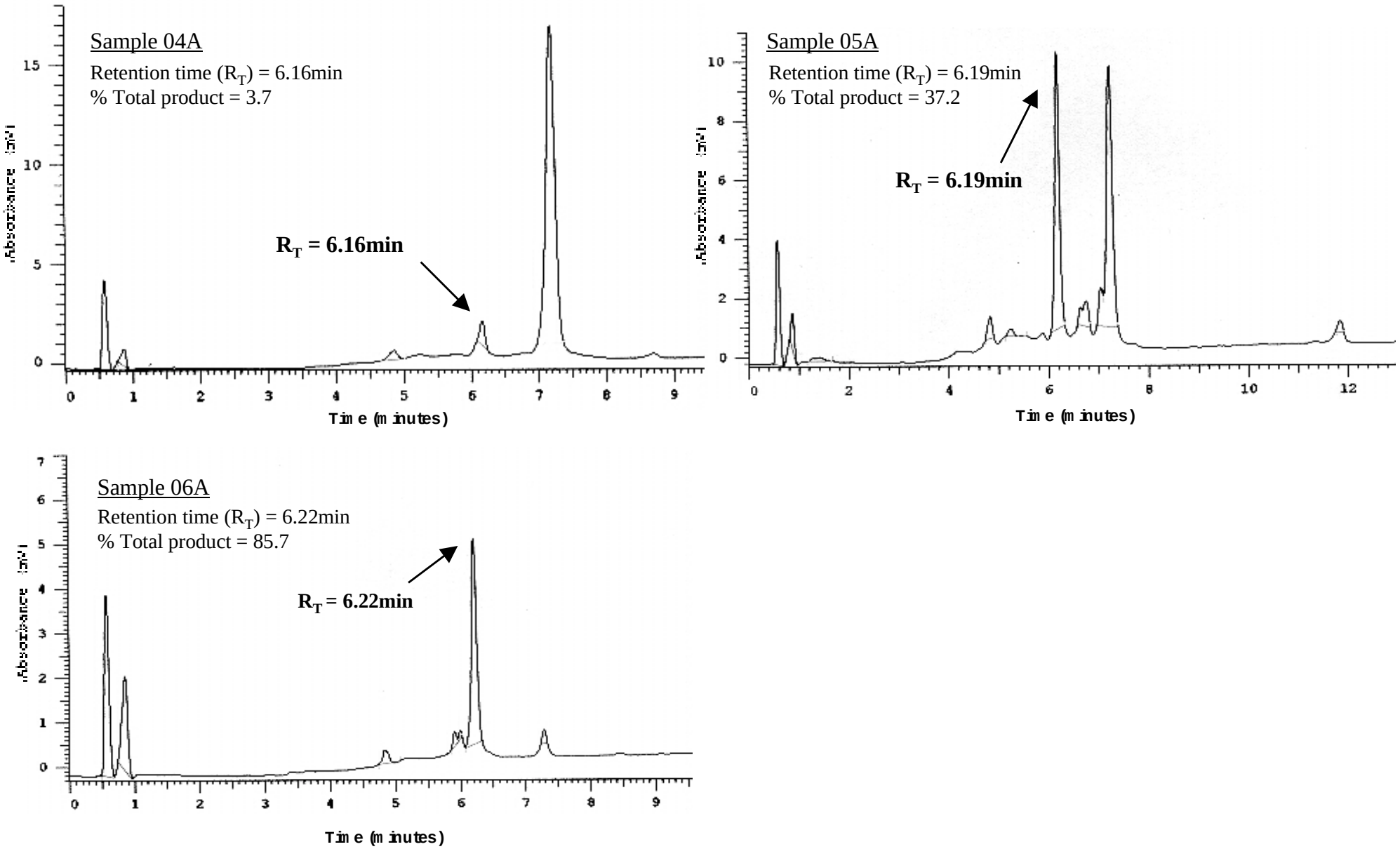

Figure 4 - Figure 4 shows HPLC data chromatograms for samples 04A-06A amplified with MPS2B (F16268/R16410). In each case, the 5.0µL injection shows one or multiple nonspecific amplicons in addition to the 143bp fragment. The average retention time for the MPS product is 6.19min and the 143bp products represent a range of per cent total DNA eluted from the column (3.7-85.7%). In each case, the peak was effectively isolated from the nonspecific products, captured and sequenced for high quality, reportable data.

Figure 5 - Sequence Data for Sample 05A

C30

HPLC

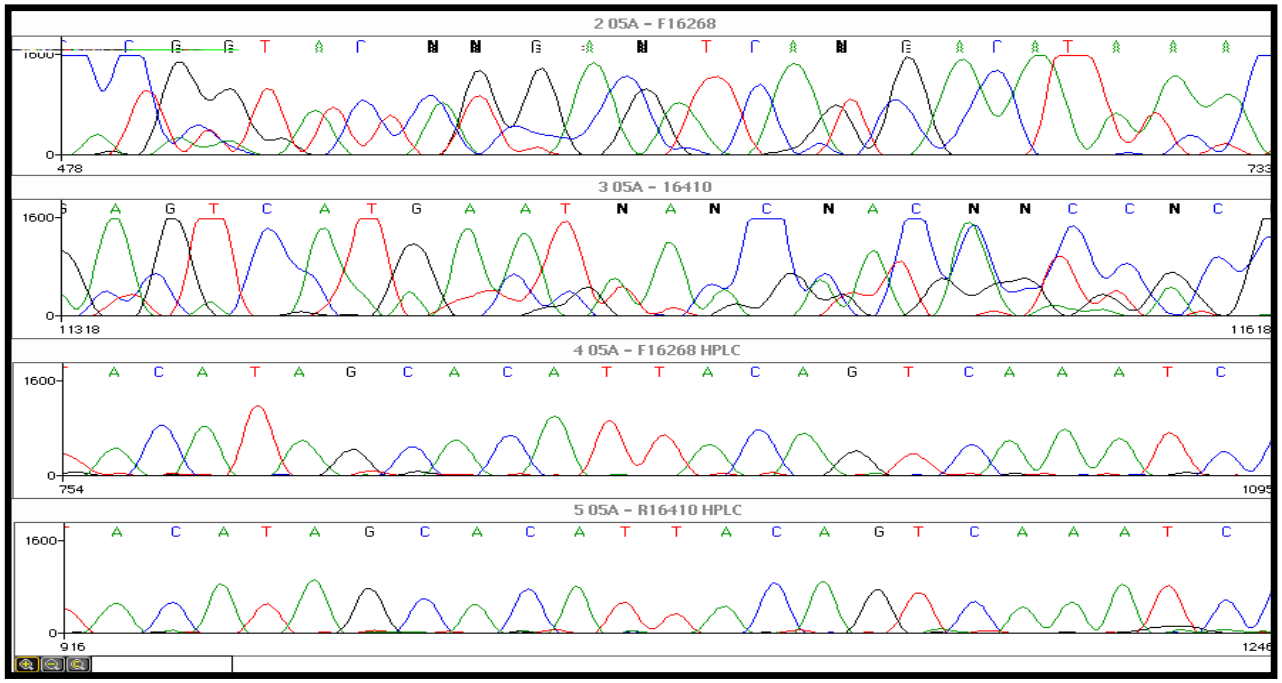

Figure 5 - Figure 5 shows sequence data electropherograms for sample 05A. The upper two panels show sequence data following purification with a Centricon-30 device that retains products based on molecular weight cutoffs. The lower two panels represent sequence data obtained following peak capture by IPRP-HPLC purification with the Transgenomic Wave System™.

Table 1 - IPRP-HPLC Sample Gradient Method for Samples 04A, 05A, and 06A

Sample Gradient

Time (min)	**% Buffer A** (0.1M TEAA)	**% Buffer B** [0.1M TEAA 25% acetonitrile]	**% Buffer C** (75% acetonitrile)	**Buffer Flow Rate** (mL/min)
0.0	65	35	0	0.75
3.0	45	55	0	0.75
10.0	35	65	0	0.75
13.5	0	0	100	0.75
16.5	0	0	100	0.75
17.5	65	35	0	0.75
19.5	65	35	0	0.75

Table 1 - Table 1 describes the sample method used for the injection and analysis of MPS2B products. In each case, 5.0µL was injected onto the column in 0.1M TEAA binding buffer (Buffer A). An increasing acetonitrile gradient (Buffer B) was introduced for DNA fragment elution at a constant flow rate of 0.75mL/min. Following sample elution, a 75% acetonitrile wash (Buffer C) was performed between sample injections to rid the column of bovine serum albumen accumulation.